



# UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE  
United States Patent and Trademark Office  
Address: COMMISSIONER FOR PATENTS  
P.O. Box 1450  
Alexandria, Virginia 22313-1450  
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
09/704,039	11/01/2000	Christopher J. Brockett	M61.12-0213	2421

7590 08/16/2004

Westman Champlin & Kelly P A  
Suite 1600 International Centre  
900 Second Avenue South  
Minneapolis, MN 55402-3319

[REDACTED] EXAMINER

SPOONER, LAMONT M

[REDACTED] ART UNIT

[REDACTED] PAPER NUMBER

2654

DATE MAILED: 08/16/2004

Please find below and/or attached an Office communication concerning this application or proceeding.

<b>Office Action Summary</b>	<b>Application No.</b>	<b>Applicant(s)</b>	
	09/704,039	BROCKETT ET AL.	
	<b>Examiner</b>	<b>Art Unit</b>	
	Lamont M Spooner	2654	

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --

#### Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 3 MONTH(S) FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If the period for reply specified above is less than thirty (30) days, a reply within the statutory minimum of thirty (30) days will be considered timely.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

#### Status

- 1) Responsive to communication(s) filed on 19 April 2004.
- 2a) This action is FINAL.                            2b) This action is non-final.
- 3) Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

#### Disposition of Claims

- 4) Claim(s) 1-24 is/are pending in the application.
- 4a) Of the above claim(s) \_\_\_\_\_ is/are withdrawn from consideration.
- 5) Claim(s) \_\_\_\_\_ is/are allowed.
- 6) Claim(s) 1-24 is/are rejected.
- 7) Claim(s) \_\_\_\_\_ is/are objected to.
- 8) Claim(s) \_\_\_\_\_ are subject to restriction and/or election requirement.

#### Application Papers

- 9) The specification is objected to by the Examiner.
- 10) The drawing(s) filed on 01 January 2000 is/are: a) accepted or b) objected to by the Examiner.  
Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).  
Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).
- 11) The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

#### Priority under 35 U.S.C. § 119

- 12) Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).
    - a) All    b) Some \* c) None of:
      1. Certified copies of the priority documents have been received.
      2. Certified copies of the priority documents have been received in Application No. \_\_\_\_\_.
      3. Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).
- \* See the attached detailed Office action for a list of the certified copies not received.

#### Attachment(s)

- |                                                                                                                          |                                                                             |
|--------------------------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------|
| 1) <input checked="" type="checkbox"/> Notice of References Cited (PTO-892)                                              | 4) <input type="checkbox"/> Interview Summary (PTO-413)                     |
| 2) <input type="checkbox"/> Notice of Draftsperson's Patent Drawing Review (PTO-948)                                     | Paper No(s)/Mail Date: _____.                                               |
| 3) <input type="checkbox"/> Information Disclosure Statement(s) (PTO-1449 or PTO/SB/08)<br>Paper No(s)/Mail Date: _____. | 5) <input type="checkbox"/> Notice of Informal Patent Application (PTO-152) |
|                                                                                                                          | 6) <input type="checkbox"/> Other: _____.                                   |

## DETAILED ACTION

### ***Response to Arguments***

1. Applicant's arguments with respect to claims 1, 9, 11, and 15 have been considered but are moot in view of the new ground(s) of rejection.

### ***Claim Rejections - 35 USC § 103***

2. The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

(a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negated by the manner in which the invention was made.

3. Claims 1-8 are rejected under 35 U.S.C. 103(a) as being unpatentable over Sproat et al., "A Stochastic Finite-State Word-Segmentation Algorithm for Chinese". Computational Linguistics, vol. 22:3, September 1996 in view of Kanevsky et al. (US Patent No. 5,835,888 Nov. 10, 1998).

Sproat et al. and Kanevsky et al. are analogous art in that they both involve information analysis within the linguistic field.

As per **claim 1**, Sproat et al. disclose a word segmentation method comprising: identifying possible segments in the sequence of characters, at least two of the possible segments overlapping each other, using a "maximum matching method... one instance of which [Sproat et al. calls] the 'greedy algorithm'" (see method described in pp. 382-384, 393-394); identifying an alternative sequence of characters for at least one of the possible segments, the alternative sequence of characters forming an alternative

segment by choosing other possible characters that may also form a segment (see "maximum matching", greedy algorithm description cited above); performing multiple syntactic analyses using the possible segments and the alternative segment, the analyses resulting in a full syntactic parse that utilizes and thereby results in a segmentation of the input sequence of characters (again please refer to the discussion of "maximum matching", greedy algorithm).

Sproat et al. does not disclose:

identifying an alternative sequence of characters for at least one of the possible segments, the alternative sequence of characters forming an alternative segment that fills the same space as one of the possible segments.

However, Kanevsky et al. teaches (Fig. 4 item 412, C.3.lines 36-43) of having alternative segments that fill the same space as one of a possible segment. Therefore, at the time of the invention, it would have been obvious to one ordinarily skilled in the art to combine Sproat et al. with Kanevsky et al. The motivation for doing so would have been to identify all possible endings in an inflected language during segmentation (C.2.lines 39-41).

As per **claims 2, 3 and 4**, Sproat et al. and Kanevsky et al. disclose a method according to claim 1, Sproat et al. further discloses:

performing multiple syntactic analyses comprises performing analyses that result in a parse containing the alternate segment (see description of "maximum matching", greedy algorithm);

As per **claim 5**, Sproat et al. and Kanevsky et al. disclose a method according to claim 1, Sproat et al. further discloses: wherein: identifying an alternative sequence of characters comprises performing inflectional morphology on a possible segment by including "morphological rules" in order to expand the dictionary used in nonstochastic lexical-knowledge-based word segmentation methods (see Sproat et al., pp. 382-389).

As per **claim 6**, Sproat et al. and Kanevsky et al. disclose a method according to claim 1, Sproat et al. further discloses:

identifying an alternative sequence of characters comprises identifying orthographic variations of a possible segment (Sproat et al., p. 384), a fact elucidated by the discussion of orthographic words in languages such as Chinese, a language in which orthographic words are written with no spaces between them, and orthographic forms are identified through use of a dictionary (Sproat et al., pp. 378-379).

As per **claim 7**, Sproat et al. and Kanevsky et al. disclose a method according to claim 6, Sproat et al. further discloses:

identifying orthographic variations comprises identifying a preferred orthographic form for the possible segment (*ibid*, via dictionary lookup).

As per **claim 8**, Sproat et al. and Kanevsky et al. disclose a method according to claim 1, Sproat et al. further discloses:

identifying orthographic variations comprises identifying orthographic variants that use a script other than the script of the characters in the possible segment by looking in Roman and other alphabets (see Sproat et al., p. 384, bottom paragraph).

Art Unit: 2654

4. Claims 9-14 are rejected under 35 U.S.C. 103(a) as being unpatentable over International Patent WO 98/08169 to Carus et al. in view of Jacquemin et al. (US Patent No. 6,101,492 Aug. 8, 2000).

Carus et al. and Jacquemin et al. are analogous art in that they both involve information analysis within the linguistic field.

As per **claim 9**, Carus et al. teach a system for identifying syntax in a string of characters from a non-segmented language, the system comprising:

a word breaker that generates a collection of words from the string of characters, the collection of words comprising at least two words that are derived in part from the same character in the string of characters (see fig. 3, items 60, 62, 64, 76; p. 6,11. 19-33; p. 29,1. 17-p. 30, 1. 9; examples: pp. 32-35), the word breaker utilizing: a lexical record set (via dictionary lookups) that is used to derive words from the collection of words by taking the words directly from the string of characters (see p. 3, 1. 23; p. 5,11. 6-18; p. 6,11. 19-33; p. 361. 28-p. 371. 10) through a database analysis module and a heuristic analysis module and;

a variants constructor that is used to derive word variants of words found in the string of characters, each word variant being added to the collection of words and having a different sequence of characters than the sequence of characters associated with the word in the string of characters from which it is derived (see previous cite, descriptions of heuristic and database analysis modules; p. 36,1. 14-p. 36, 1. 10; fig. 5, step 100; fig. 6); and

a syntax parser that performs a syntactic analysis using the collection of words produced by the word breaker to produce a syntax parse; the process outlined in Carus et al. therein comprising a syntax parse which identifies syntactic units of a given corpus per applicant's claim and specification (see fig. 2; p. 9,11. 22-36; p. 10,11. 1-9).

Carus does not disclose:

a variants constructor that is used to derive word variants of words found in the string of characters where the word variants are not found in the string of characters.

However, Jacquemin et al. teaches (C.9.lines 30-35) deriving word variants to a word containing variants of the word not including in the string containing the word. Therefore, at the time of the invention, it would have been obvious to one ordinarily skilled in the art to combine Carus et al. with Jacquemin et al. The motivation for doing so would have been to provide all possible derivatives of a word in order for determining the syntactical structure of a word for further processing (C.4.lines 1-8).

As per **claim 10**, Carus et al. and Jacquemin et al. teach a system according to claim 9, Carus et al. further discloses:

the variants constructor comprises inflectional morphology rules (see previous site under variants constructor, claim 9).

As per **claim 11**, Carus et al. and Jacquemin et al. disclose the teachings of claim 10, on which claim 11 depends (see cite in claim 10 rejection above).

Carus et al. fail to explicitly teach inflectional morphology rules that are capable of identifying a word's lemma from its inflectional form in the string of characters.

Art Unit: 2654

However, Jacquemin et al. disclose inflectional morphology rules that can identify a word's stem from inflectional morphology rules (C.4.lines 57-61). At the time of the invention, it would have been obvious for one skilled in the art to modify Carus et al. to include inflectional morphology rules capable of identifying a word's stem from an inflectional form because Jacquemin et al. teach the use of inflectional morphology (C.4.lines 57-61) rules in order to better disambiguate information, as for use in further syntactical analysis (C.5.lines 19-32).

As per **claim 12**, Carus et al. and Jacquemin et al. teach a system according to claim 9, Carus et al. further discloses:

the variants constructor comprises an orthographic variants structure that indicates the orthographic variants of words found in the string of characters (see previous cites in claim 9 regarding the lexical record set and variants constructor), a dictionary being used to check character orthography.

As per **claim 13**, Carus et al. and Jacquemin et al. teach a system according to claim 9, Carus et al. further discloses:

at least one word variant has a different number of characters than the word from which it is derived, the variants constructor disclosed in Carus necessarily creating a variant with a different number of characters (see previous cite in claim 9 regarding variants constructor).

As per **claim 14**, Carus et al. and Jacquemin et al. teach a system according to claim 9, Carus et al. further discloses:

at least one word variant includes a character that is not present in the string of characters, the variants constructor disclosed in Carus necessarily creating a variant with alternate characters (see previous cite in claim 9 regarding variants constructor).

5. Claims 15-24 are rejected under 35 U.S.C. 103(a) as being unpatentable over Carus et al. in view of Jacquemin et al., further in view of Sproat et al, and further in view of Kanevsky et al.

As per claims 15-24, Carus et al., Sproat et al., Jacquemin et al. and Kanevsky et al. disclose methods according to claims 1-14.

Carus et al., Jacquemin et al. Kanevsky et al. and Sproat et al. do not specifically disclose a computer system and a computer-readable medium to execute said methods.

At the time of invention, it would have been obvious to one skilled in the art to modify Carus et al., Jacquemin et al., Kanevsky et al. and Sproat et al. to include a computer system to execute the disclosed methods and a computer-readable medium coupled to said system; such modification is implicit in the assumptions and spirit of both inventions. Such a modification would have been obvious in order to execute said methods, update the method and executable code, provide read-write storage for the executable code and data, evaluate its performance, and incorporate the system and media into a product.

***Conclusion***

6. Any inquiry concerning this communication or earlier communications from the examiner should be directed to Lamont M Spooner whose telephone number is 703/305-8661. The examiner can normally be reached on 8:00 AM - 5:00 PM.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Nguyen Vo can be reached on 703/308-6728. The fax phone number for the organization where this application or proceeding is assigned is 703-872-9306.

Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see <http://pair-direct.uspto.gov>. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free).

Ims  
08/09/2004



PATRICK N. EDOUARD  
PRIMARY EXAMINER